DOCUMENT RESUME

ED 253 543 TM 840 747

AUTHOR Worthen, Blaine R.; Sanders, James R.
TITLE Content Specialization and Educational Evaluation: A Necessary Marriage? Paper #14 in Occasional Paper Series.
INSTITUTION Western Michigan Univ., Kalamazoo. Evaluation Center.
PUB DATE Sep 84
NOTE 45p.
AVAILABLE FROM Evaluation Center, Western Michigan University, Kalamazoo, MI 49008 ($3.50).
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Educational Assessment; Elementary Secondary Education; *Evaluation; *Evaluation Methods; *Evaluators; Higher Education; *Intellectual Disciplines; *Professional Training; Specialization

ABSTRACT

Should an educational evaluator be trained as a content specialist, an evaluation specialist, or some combination of the two? A content specialist has demonstratable expertise in the field of study being evaluated, but not training in evaluation methods; the evaluation specialist has expertise in evaluation methods but may know little about the subject area being evaluated. The merits of each type of evaluator are analyzed in six areas: (1) the difficulty and uniqueness of the subject content; (2) reference groups and impartiality; (3) the roles and tasks of evaluation; (4) the evaluator's scope of work; (5) implications for evaluator training; and (6) professional status and rewards. It is concluded that evaluation specialists are the best choice to evaluate most educational enterprises. Some attention is given to the problem of how professional evaluators may effectively manage those portions of the evaluation needing the content expertise they do not have. Content specialization plays an important role in educational evaluation, but it is neither necessary nor desirable in the training of educational evaluators. (BS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED253543

Content Specialization and Educational Evaluation: A Necessary Marriage?[1]

by

Blaine R. Worthen

James R. Sanders

September 1984

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

X Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Sanders & B. Worthen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EVALUATION CENTER

COLLEGE OF EDUCATION
WESTERN MICHIGAN UNIVERSITY
KALAMAZOO, MICHIGAN 49008-3899

TM 840 747

Paper #14
in
Occasional Paper Series

BEST COPY AVAILABLE

WESTERN MICHIGAN UNIVERSITY
KALAMAZOO, MICHIGAN 49008-3899

EVALUATION REPORT SERIES

The Evaluation Report Series is published by the Evaluation Center, College of Education, Western Michigan University. It is designed to provide models of evaluation design or reporting techniques. Selected reports will be provided as part of the Center's dissemination activities. Reports are reprinted with permission of the original publisher.

For information about previous publications in the series, see the inside of the back cover of this document.

All correspondence in reference to the Evaluation Report Series should be addressed to:

Information Officer
Evaluation Center
College of Education
Western Michigan University
Kalamazoo, Michigan 49008-3899

EVALUATION CENTER

Daniel L. Stufflebeam
Director

Mary Anne Bunda

James R. Sanders

ABOUT THE CENTER

The Evaluation Center is committed to advancing the science and practice of evaluation. To serve its purpose the Center has developed components in research, development, instruction, dissemination, and service. Over the past ten years the staff of the Center has had significant experience in these areas working with school districts, state departments of education, foundations, and federal agencies.

The *Research Component* is designed to develop and test evaluation theory. Particular thrusts are testing the effects of school-based evaluation offices on school practice, and determining the special evaluation skill and knowledge profiles required by various groups in education. The *Development Component* is designed to work on ways and means of implementing evaluation theory. Included are techniques for meeting special evaluation objectives, new data collection instruments, and special procedures for communicating evaluative findings and relating them to practical problems. The *Instructional Component* is designed to provide in-service training in evaluation to various local, state, and national groups. While the Center is not a credit-granting agency, it collaborates with instructional units throughout Western Michigan University for the development and implementation of pre-service evaluation curricula. Emphasis in this work is at the doctors and masters levels. The Center serves as an instructional laboratory wherein students majoring and minoring in evaluation can gain practical evaluation experience. The Center is also committed to the development of novel and needed evaluation training materials and strategies. The *Dissemination Component* is designed to communicate research findings, new ideas, and materials to the education profession. *The Service Component* is designed to assist agencies in meeting their evaluation needs through two primary services. These are the design and implementation of evaluation studies and the development and installation of agency evaluation systems.

The staff of the Center has been involved in conceptualizing the CIPP Evaluation Model, researching the Item Sampling technique, developing the Advocate Team Technique, developing over 100 standardized tests, designing and conducting many in-service evaluation programs in evaluation. Other accomplishments include authoring the PDK book, EDUCATIONAL EVALUATION AND DECISION MAKING, and numerous journal articles and papers, designing and conducting many evaluation projects, and installing evaluation systems in schools, government agencies, and research agencies. Overall, the Center staff has directed externally funded projects in evaluation totaling more than four million dollars.

CONTENT SPECIALIZATION AND EDUCATIONAL EVALUATION:

A NECESSARY MARRIAGE?[1]

Blaine R. Worthen
Utah State University

James R. Sanders
Western Michigan University

September 1984

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Sanders/B. Worthen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 840 747

Despite the present economic difficulties facing all levels of our nation's educational enterprise, a significant portion of education's resources continue to be used for evaluative activities. As a result, presumably competent evaluation personnel are sought after to evaluate activities in nearly every area of education and its cognate fields. Persons are being assigned to evaluate social action programs, experiential learning, dental education, bilingual education, computer-assisted instruction, aesthetic education, and new curricula in geology, mathematics, and foreign languages, to name only a few.

Regardless of the type of educational program to be evaluated, there are some tasks that are common (e.g., identifying the evaluative questions to be answered by the evaluation; identifying and selecting sources of relevant information; collecting, interpreting and reporting information). Persons trained explicitly as evaluation specialists can reasonably be expected to have a command of these and other methods and techniques of educational inquiry necessary in evaluation. But is this enough, given the enormous diversity in the content with which evaluators are asked to deal? Does the training of educational evaluators prepare them to deal effectively with such a wide range of phenomena? Or does quality evaluation demand that the evaluator have some substantive training in the content of that being evaluated? For example, must an evaluator be trained in guidance techniques and theories before producing ar accurate, useful evaluation of a guidance and counseling program?

These questions are more than academic, for they are frequently posed to evaluation specialists who are hired or assigned to conduct evaluations in areas where they hold no academic credentials. Such questions could be discounted as only new manifestations of the resistance and inhospitality evaluators have come to expect, since the field of education is still largely unaccustomed to (and unenthused about) serious introspection into the effectiveness of its practices. But such an interpretation seems grossly unfair and insensitive to the real issue which lies behind the questions. It is probable that most such queries originate from educators' sincere desires for the best possible evaluation, coupled with genuine concern that someone untrained in the content of that which is evaluated simply cannot do the best job, even if they possess impeccable methodological credentials. We would nominate concern, not hostility, as the author of the statement, "I respect your training as an evaluator, but I don't see how you can evaluate my language program adequately, since you don't know much about linguistics." Which leads us back to our two original questions, slightly rephrased. Can evaluators with no real expertise in reading really judge the worth of a reading program? Or is it preferable to hire reading experts and provide them with on-the-job training in evaluation skills? Or is it realistic to look for both content and evaluation expertise in the same person?

Answers to these questions are impossible without first clarifying what is meant by the term "content specialization." Evaluation has been defined previously (Worthen & Sanders, 1973; Joint Committee, 1981) and

we will not repeat that exercise here, although we will return later to point out how answers to the questions posed above might differ somewhat depending on the evaluation role or approach in use.

Content Specialization: A Definition

The term "content specialization" might be used to refer merely to concentration of one's efforts in an activity or field of study. That arid definition is insufficient for the present purposes without adding that the concentration leads to wide knowledge about and competence to work in the field. Therefore, content specialization is used throughout this paper with a more qualitative tone to refer to demonstrable expertise in the matter dealt with in a field of study or activity. Some elasticity must be retained in applying the term because of the varied evaluation contexts described earlier. For example, in some cases understanding of theories or detailed knowledge of subject matter might be the most relevant expertise. In others, years of practical experience might be the touchstone, with the wisdom of the expert practitioner elevated above the knowledge of the scholar. The common ingredient in all uses of the term is that the principal participants (e.g., director, staff) in any activity being evaluated would recognize and credit the expertise of the "content specialist."

The remainder of this paper deals with three major topics: (1) some of the possible interrelationships between content specialization and evaluation; (2) a series of basic points which must be considered in attempting to answer the questions posed earlier; and (3) considerations

in implementing the answer we propose to those questions. Another related topic--namely, how the content area of the object being evaluated influences the evaluation methodology used--is beyond the scope of this paper.

Three Alternative Profiles of Expertise

If one accepts the premise that no one person can have all the expertise needed to conduct an evaluation adequately, then it would be reasonable to argue that the best way to bring necessary evaluation and content specialization into play would be to use a team of experts in conducting an educational evaluation. This alternative would allow for the right *ad hoc* configuration of experts (evaluation specialist, content specialist, computer systems analyst, etc.) to be drawn together for a particular evaluation. When conducting an evaluation, it is delightful to have one member of the team who is expert in the relevant content area. When other team members are baffled by jargon and verbal clutter, this content specialist should be able to cut through quickly to lay bare the skeletal theories and untested assertions at the heart of an issue.

This alternative, while desirable when organizing to conduct an evaluation, misses the issue under consideration here, however. The point of our discussion is to determine whether it is best to prepare the individual evaluator in education as a content specialist, an evaluation specialist, or some combination of the two. Many educational evaluations are conducted by only one evaluator, and such an individual cannot be prepared as a team. The skill of pulling together an appropriate team of

specialists for an evaluation is a skill that the professional evaluator should possess, as will be discussed later in this paper. This discussion focuses, however, on training for the individual evaluator, and here there are three obvious alternatives for bringing necessary evaluation and content expertise into play in educational evaluations.

Content Specialist. The first alternative is to entrust the evaluation to someone trained as a content specialist in the areas most relevant to the entity being evaluated. The content specialist would need to learn as much about evaluation methods and techniques as possible and depend on expert methodologists where the evaluation demands knowledge beyond that held by the substantive specialist.

Professional Evaluator. In the second alternative, the evaluation is conducted by a specialist trained specifically in methods and techniques of educational inquiry necessary in evaluation. The professional evaluator would need to learn as much as would be useful about the content area and depend on content specialists whenever the evaluation required detailed or extensive knowledge about the subject area. The professional evaluator might work concurrently or sequentially on evaluations in many different content areas, claiming expertise only in evaluation methodology and making no pretense of being expert in the content.[2]

Content-based Evaluator. This is really a combination of the two previous alternatives. Here the evaluator is either (1) a professional evaluator who has worked in the same content area over a long period and gained sufficient expertise in it to be viewed as a content specialist,

(2) a content specialist who has been assigned to evaluation roles for so long as to master the principal methods and techniques of educational evaluation, or (3) a person who holds academic credentials in both areas.

Before making judgments about the relative desirability of these three alternatives, it is necessary to examine several factors which complicate the choice.

Basic Considerations in Electing Content Specialization or Evaluation Expertise

The issue of subject matter specialization for evaluators has been touched on previously. In analyzing the role of the evaluator, Stufflebeam and other evaluation scholars concluded that the role of subject matter specialist "...must be included if evaluation is to serve decision making. However, it does not seem to be a role that the evaluation specialist can often assume." (Stufflebeam, et al., 1971, p. 294). Scriven at least partially supported this view in his statement that "...the evaluator, while a professional in his own field, is usually not a professional in the field relevant to the curriculum being reformed or, if he is, he is not committed to the particular development being undertaken" (Scriven, 1973, p. 65).[3] Conversely, Cronbach et al. (1980) proposed that disciplinary preparation be one of four parts of an "idealized" doctoral training program in evaluation. The Joint Committee on Standards for Educational Evaluation (1981) further noted, "Evaluators are credible to the extent that they exhibit the training, technical competence, substantive knowledge, experience, integrity, public

relations skills, and other characteristics considered necessary by the client and other users of the evaluation reports. (Joint Committee, 1981, p. 24, italics added). Davis, Scriven and Thomas (1981) also advised educators that an evaluation study can best be defended from attack by having it conducted by persons who are viewed as credible experts by the major parties involved. To better understand issues underlying these varying assertions, let us turn our attention directly to six factors which impinge on them more or less directly.

1. Difficulty and Uniqueness of the Content

It may be truistic to state that concern about the professional evaluator's grasp of content is not very relevant where the content is neither difficult nor unique and therefore can be easily assimilated and understood. But this point is generally lost on the persons who express the concern most. The field of education obviously has its complex theories and difficult content (largely drawn from other fields). However, the knowledge base in education is not as complicated as it sometimes is made to seem to the outsider who is forced to sort through many private meanings in the language of the educationist to reach understanding. Most theories in education are essentially primitive and most educational practices can be easily comprehended if they are clearly described.[4] It is probably neither arrogance nor criticism which has led some social scientists to privately aver that they can digest even the most complex educational theories in an afternoon.

The opinion of the social scientist is less important here, however, than the magnitude of the task confronting educational evaluators when they stray onto foreign ground. Often the task does not appear overwhelming. It should not take long to learn all any evaluator would need to know about computer-assisted instruction (CAI)--the basic rationale, results of previous CAI research, the particular situation in the instance at hand, and how it operates in the real settings observed. The evaluator may not know enough of the intricacies involved in setting up the logistics of CAI to enable replication of the activity, but that seems unlikely to be the best criterion for judging the adequacy of the program in any event.

The notion that simplicity[5] should be less valued than complexity is an egregious one that seems to influence much thinking in the field of education (and other fields as well). The earlier point about complex transactions and outcome notwithstanding, there are many things about education which are refreshingly simple. Neither that nor the fact that we have not yet reached a full understanding of the things in education which are complex should embarrass educationists. Yet, it must be fear of having our theories and practices demeaned as simple-minded (in the most pejorative sense of that word) which causes many educational specialists to contend that their work is too complex and filled with subtle nuances to be readily understood by outsiders. The alternative explanations for this tendency are less complimentary.[6]

The message here is not that all aspects of educational theory and practice are uncomplicated and easy to grasp; that stance would be absurd on its face. The point is only that much of the work being conducted in the field involves subject matter or practices that are neither difficult nor unique and, in these areas, it seems irrelevant (if not foolish) to argue the need for content specialists to serve as evaluators.

2. Reference Groups and Impartiality

Intuitively, it seems that there is a difference between being a judge and being a processor of judgments made by others; between using one's personal values to reach judgments of worth and using the collective values of appropriate reference groups to identify criteria to serve as a basis for such judgments. And there are obvious values involved even in determining which reference groups are appropriate. The importance of these areas in educational evaluation has been well documented by the dialogue they have created in the literature. Scriven (1973) has argued that there is no evaluation without judgment and the evaluator is best qualified to judge. Stufflebeam, et al., (1971) agreed that an evaluation depends on judgments of the worth of alternatives, although they saw the evaluator as more clarifier and arbitrator than final judge. In their view, the evaluator would help identify different value positions of many reference groups which might influence the final decision and help the "decision group" understand the risks of attending more to one position than another. Cronbach (1982; Cronbach et al., 1980) likewise described evaluation as basically a democratic process.

Stake (1973; 1975a) opted for inclusion of multiple reference groups as "judges" or at least sources of criteria, including as a minimum societal representatives, teachers, parents, students, and subject-matter experts. He stated that:

> "Evaluators will seek out and record the opinions of persons of special qualification. These opinions, though subjective, can be very useful and can be gathered objectively, independent of the solicitor's opinions. A responsibility for processing judgments is much more acceptable to the evaluation specialist than one for rendering judgments himself. (Stake, 1973, p. 111).

Stake also has reminded us that judgment consists of assigning weights to standards. "Rational judgment in educational evaluation is a decision as to how much to pay attention to the standards of each reference group (point of view) in deciding whether or not to take some administrative action." (Stake, 1973, p. 122).

This leads us to a central concern. What influence do the formal training and experience of individuals (i.e., the reference group or groups to which they belong) have on their inclination or ability to identify all the reference groups from which values, criteria and opinions should be sought in a particular evaluation? The question can be particularized to the present discussion of evaluation and content specialization. Are professional evaluators disposed to be satisfied with their own judgments of the worth of a program? Are content specialists less inclined to seek input from other reference groups because they view themselves as expert in the area?

These are largely empirical questions, but until they are answered that way, we will venture some "best guesses", and these guesses stem from a concern about what happens when the evaluator's personal values intrude in unidentified ways into value judgments about that which is evaluated.[7] This concern should not be misconstrued as an argument against making value judgments. It is only an argument against basing those judgments on personal, private, and perhaps idiosyncratic values of the evaluator in ways that preclude their being identified or sorted out either by the evaluator or the consumer of the evaluation reports. To say that value questions are the sine qua non of evaluation (see Glass & Worthen, 1971) is not to say that it is the evaluator's values which should be used to resolve those questions. When Stufflebeam and his colleagues stated that, "The evaluator is the supplier of knowledge. He never supplies the values with which that knowledge is used," (Stufflebeam, et al., 1971, p. 117), they were arguing that evaluators never supply their personal values, but instead that they should help identify various value positions and look for optimal combinations of those values to present to the decision group. The Joint Committee on Standards for Educational Evaluation (1981) identified the following value bases that could be used by the evaluator: "...project objectives, procedural specifications, laws and regulations, institutional goals, democratic ideals, performance by a comparison group, assessed needs of a consumer group, expected performance of the sample group, and reported judgments by various reference groups." (Joint Committee, 1981, p. 32).

The intrusion of the evaluator's personal values seems more probable when the evaluator is a content specialist in the area evaluated. Here the content expert is on home ground and can bring personal expertise into play directly, and that very fact has a great potential for blinding the expert to the need to elicit opinions and judgments from other reference groups, perhaps including others with expertise in the content area.[8] Less so for professional evaluators. They certainly carry their own prejudices, but they should prove less damaging here. Unless they suffer severely from exaggerated self-worth, evaluation specialists are virtually forced to a recognition of their own naivete and need for help from outside reference groups whenever asked to conduct an evaluation involving difficult or unique content outside of their training and experience.

It would seem, at least on the surface, that content specialists evaluating activities in their own particular area of expertise might well have considerable difficulty retaining their independence and impartiality in the face of their identification and career ties in that field. Scriven seemed to have this in mind when he said:

> "Evaluators...are handicapped so long as they are less than fully familiar with the subject matter being restructured, and less than fully sympathetic with the aims of the creative group. Yet once they become identified with those aims, emotionally as well as economically, they lose something of great importance to an objective evaluation--their independence." (Scriven, 1973, p.66).

A biologist called to evaluate a curriculum developed by biologists is prone to look forward to future professional associations with the curriculum developers, perhaps for the remainder of a long career. The influence that acceptance by peers has on later success can hardly help creating at least a threat to impartiality, whether realized or not. This issue was touched upon in the Joint committee "Conflict of Interest" standard (1981, p.70).

It would be folly to suggest that being a professional evaluation specialist is any guarantee against loss of independence, since both the biologist and the professional evaluator can be compromised through economic or emotional identification with the curriculum, but the evaluator would seem to have less opportunity for conflict of interest than would the content specialist.

Weiss, in quoting an informant from her study for the National Institute of Mental Health, introduced two related points:

> "As one psychiatrist noted, there is a basic difference in stance between practitioners and evaluators: 'Practitioners have to believe in what they are doing; evaluators have to doubt'.
>
> "This difference in professional orientation also can be seen in the individual's orientation to the project. Practitioners are committed to a project; they invest enormous amounts of time and energy and their professional reputations in its success. Evaluators are committed to the acquisition of knowledge, and their careers are dependent on producing competent research whether the project succeeds or fails; thus, they are sometimes viewed as unsympathetic and perhaps basically critical of the project to which others are devoting their lives." (Weiss, 1973, p. 52).

To assume that content specialists are caught up in commitment to an activity as much as practitioners directly involved in the project would be extending a point too far. It seems equally risky, however, to assume that involvement with the content does not carry a greater potential for bias than where no such involvement exists.

Weiss' description of the negative perceptions which can result from different basic orientations has also been noted by Scriven. He commented on the effect the "Doubting Thomas" nature of evaluators could have on curriculum developers:

> "Professional evaluators may simply exude a kind of skeptical spirit that dampens the creative fires of a productive group. They may be sympathetic but impose such crushing demands on operational formulation of goals by the group as to divert too much time to an essentially secondary activity." (Scriven, 1973, p. 66).

Uncontrolled skepticism obviously can be detrimental, as can be almost any excess, including uncontrolled enthusiasm. But skepticism is not synonymous with scorn, and a circumspect, well-motivated skepticism which challenges easy assumptions and conventional wisdom that find their way into many educational activities may be invaluable to the ultimate success of those activities. If that statement causes administrators and curriculum developers to blanch, they have misunderstood our point, for it is one of affirmation. Many educational materials or processes are nurtured to full term in a sheltered environment, only to find that they cannot survive in the world of reality into which they are delivered. If the professional evaluator exudes the kind of skeptical spirit which prevents that type

of disaster, it would seem to foster a healthy and functional type of self-scrutiny indeed. It is probably unreasonable to hope that content specialists who are immersed in one way or another in the content being evaluated will have as great a tendency to raise the difficult questions which skepticism prompts.

3. Evaluation Roles and Tasks

There is an obvious omission in the discussion so far, and that is the effect of different roles and conceptions of evaluation on the points which have been made. For example, to the extent concerns about cooptation and bias are less relevant for formative than for summative evaluation, the preceding discussion could be viewed as more applicable to the latter. But even in internal formative evaluation (let alone external, goal-free formative evaluation), it seems unwise to ignore the loss of objectivity which occurs when an evaluator becomes coopted into the reference group being served and begins to accept their assumptions and procedures without question. To the extent that being a content specialist in the area evaluated may increase one's susceptibility to such loss of independence, the previous discussion is relevant to formative evaluation as well.

When one examines different conceptions of evaluation, the strain on the previous arguments becomes more apparent. For example, if an accreditation or professional judgment evaluation approach is chosen and the evaluator is nominated as the person who judges the merit of a set of curriculum materials, then it seems reasonable to assume that

the evaluator needs considerable knowledge about the content of those materials. Conversely, if evaluation is viewed as seeking, processing, and portraying judgments of multiple reference groups (including content specialists), then there seems little point in requiring the evaluator to have in-depth knowledge about the phenomena being evaluated.

Much more could be said about roles and approaches to evaluation in this connection, but it seems more important to turn to the more salient matter of what evaluators do when they evaluate.

Perhaps the most serious test of the proposition that evaluators do (or do not) need to be well-versed in the content of that which they evaluate is to examine the tasks that are essential in conducting an educational evaluation. Sorting essential competencies into those which require knowledge of substantive knowledge about the phenomenon versus those requiring specialized knowledge of evaluation methodology is enlightening. In one such analysis, Worthen (1974) used 25 evaluation tasks from a synthesis of empirical and conceptual research aimed at identifying essential tasks and competencies in educational research and evaluation.[10] These tasks are arrayed in Table 1 along with a rating of the probable need for training and experience in both evaluation and the content area. Crude categorization results in 13 tasks which an evaluation specialist would seem best able to perform, two which favor the content specialist and 10 which seem equally suited to either.[11] One might quarrel with where checkmarks were placed, but it seems unlikely that any reasonable rearrangement would result in a

significant change in the balance reflected here. The obvious inference to be drawn from this analysis is that content-based evaluators (those persons described earlier as possessing expertise both in evaluation and the content area) would be able to perform all the tasks. If such a person were not available (which seems probable for reasons discussed later), then the professional evaluator appears able to conduct far more of the essential evaluation activities than is true for the content specialist.

The tasks listed in Table 1 admittedly reflect an empirical, behavioral bias; their only defense is that they are synthesized from an analysis of high-quality evaluation work and from evaluation tasks nominated as essential by 14 leading evaluators. Further, they are quite parallel to *a priori* conceptions of evaluation theorists. For example, Stufflebeam (1973) has listed 22 steps in designing educational evaluations which he offers as potentially useful in defining the role of the evaluation specialist. Of these, at least 20 seem more appropriately the province of the professional evaluator than the content specialist, and the other two are debatable. Although Stufflebeam's design steps require frequent interactions between the decision maker and the evaluator, there is no claim that the latter must be a specialist in the content of what is evaluated. This was consistent with the PDK Study Committee (Stufflebeam, *et al.*, 1971), when they posited four categories of knowledge needed by the evaluation specialist and subsumed all 22 of Stufflebeam's design steps under "Knowledge required in evaluative work." Only one of the other three

categories includes knowledge from areas outside of evaluation, and here the only three areas suggested as useful to the evaluation specialist were general systems theory, economics, and political science. The Study Committee summed up their position with the statement that "...failure to include subject expertise in their description of evaluation activities or roles suggests that this role, like the decision-maker role, is involved in evaluation but in a way that differentiates it from the role of the evaluation specialist." (Stufflebeam, et al., 1971, p. 296).

In a somewhat different type of listing, Scriven (1974b) proposed a checklist for use in evaluating educational products, proposals focused on products, producers of products, and the like. That checklist and the broad applicability claimed for its criteria seem to rest on an assumption that an evaluator need not be a specialist in the content of the product to use it effectively. The information required to apply each criterion is straightforward and uncomplicated[12] and does not call for elaborate details or complex rationales for the product which could only be supplied by a content specialists. Scriven has noted that the evaluator must call in independent experts to judge the educational significance of performance data, reinforcing the notion that he sees the professional evaluator as the prime judge, calling on content experts only as needed. Content specialist untrained in evaluation methods could hardly hope to generate adequate information on several of the prcposed criteria, such as those related to field tests, performance/causation, and performance/statistical significance.

In an integrative review of earlier studies and discussions of evaluation competencies, Sanders (1981) identified 11 categories of competence or ability,[13] including the ability to:

1. describe fully the entity being evaluated;
2. describe the pertinent context of the evaluation;
3. conceptualize appropriate purposes and frameworks for the evaluation;
4. identify and select appropriate information needs and sources;
5. identify, select, and apply appropriate techniques and procedures for information collection, processing and analysis;
6. determine the value or merit in the entity being evaluated;
7. communicate evaluation plans and results effectively;
8. manage the evaluation;
9. maintain ethical evaluation standards;
10. adjust for external factors that affect the evaluation; and
11. evaluate the evaluation.

Although at a more general level than the specific competencies listed in earlier works and reflecting a greater emphasis on conceptual competence, even a conservative reading of these general activities suggests that the evaluation specialist will be better prepared to perform a majority of them effectively than would the content specialist.

Worthen (1983) attempted to assess the relevance of the 25 evaluation tasks shown in Table 1 to three evaluation studies which had received recognition by national professional associations or governmental bodies as examples of excellent evaluation work. He found that almost all of the tasks were important in conducting those three exemplary evaluation studies. Exceptions were tasks relating to the general area of goals and objectives (tasks 10-12 in Table 1), which were relatively less important in those three studies, suggesting that newer conceptions of evaluation may have weakened the centrality of the goal-directed evaluation approach and its frequent insistence that goals be translated into specific behavioral objectives. In addition, Worthen identified in these three exemplary evaluations five additional tasks important to the conduct of those studies; those tasks are presented in Table 2, categorized as to whether the task would be best suited to the evaluation specialist, the content specialist, or either.

This analysis again suggests that persons trained as evaluation specialists will be better prepared to conduct high-quality evaluation studies than one might expect for content specialists, as judged by tasks found to be important in exemplary evaluations.

4. The Evaluator's Scope of Work

Demands are very different for the evaluator who is employed exclusively on a single project or in one content area than for the more typical evaluator who is employed to work (either concurrently or

sequentially) in many content areas and on many different evaluations.[14] On long-term projects where the disciplinary base and methods are fixed, it is possible to assign the evaluation to a content specialist or content-based evaluator, and all the previous discussion in this paper would apply. More typically, however, the evaluation is undertaken at the behest of a client, leaving the evaluators little autonomy to pursue inquiry within any particular paradigm of their choice or within a specific subject in which they have substantive expertise. In Guttentag's terms, "The evaluation researcher does not formulate his own hypothesis. When he investigates, and how he conceptualizes what he investigates, is given to him by the program goals" (1973, p. 61). This necessity to work across content areas and on many evaluations has obvious implications for the training of evaluators which will be discussed in the next section. It is sufficient here to label as absurd the idea that any evaluator who works in multiple content areas could be so broadly knowledgeable as to be credible as a content specialist in all or even a majority of those areas.

5. Implications for Training

Some prominent evaluators have suggested, at least verbally, that training evaluation specialists is less promising than looking to the content areas for persons who are both in tune with their own field of study and introspective and capable enough to be able to evaluate their own activities. In a Utopian world, that proposal would have greater merit than in a world where time and opportunity to learn about both

evaluation methods and the content are real constraints. Content specialists with no formal training in the essential inquiry methods outlined earlier (e.g., design, psychometrics, statistics) are simply unprepared to do a majority of the evaluation tasks. Trying to provide sufficient training in even those few methods would be an enormous task. Most fulltime graduate students have their hands full mastering the bare methodological essentials in their two or three years of formal study. It would be infinitely more difficult to attain the same ends in an inservice training context where training was a part-time activity and training opportunities in the necessary methodology were sporadic at best.

Problems in training professional evaluators are another matter. Here training will probably be strongest in the essential evaluation tasks such as those listed in Table 1. Although Cronbach and others (1980) have proposed disciplinary preparation as part of an "idealized" doctoral training program, a training problem emerges here, because of the lack of autonomy the evaluator has in specifying areas for study. This problem is examined in the extended quotations below.

> "That the educational researcher can afford to pursue inquiry within one paradigm and the evaluator cannot is one of many consequences of the autonomy of inquiry. When one is free to define his own problems for solution (as the researcher is), he seldom asks a question that takes him outside of the discipline in which he was trained. Psychologists pose questions that can be solved by the methods of psychology, as do sociologists, economists, and other scientists, each to his own. The seeds of the answer to a research question are planted with the question. The curriculum evaluator enjoys less freedom in the definition of the questions he must answer. Hence, the answers are not as likely to be found by use a stereotyped methodology. Typically, then, the evaluator finds it necessary to employ a wider range of inquiry perspectives and techniques to deal with questions that do not have predestined answers." (Glass & Worthen 1971, p. 161).

> "..to the extent that the training of evaluators touches on the traditional disciplines at all, it is best that several disciplines be sampled. The trainee should appreciate the view of education afforded by each of the socially-relevant disciplines. In this way, the evaluator can become sensitive to the wide range of phenomena to which he must attend if he is to properly assess the worth of an educational program." (Glass & Worthen, 1972, p. 93).

If one accepts the assertion that evaluators should be educated broadly in the inquiry methods of several disciplines, it should be apparent that they will have little opportunity for preparation in the substantive content of any discipline. Thorough preparation in a discipline may be gained in part during a graduate program, but most often it will develop in a mature form only after many years of professional practice. Doctoral students preparing themselves for a career in evaluation leave their programs with their basic "evaluation capabilities" intact. Not so for content specialists, who may be trying to add the first layer of methodological expertise necessary to establish capability as an evaluator. Asking evaluation methodologists to extend their knowledge to encompass the methods of various disciplines is to challenge them to the utmost. Asking content specialists to do the same, ignoring basic deficits in methodology, is probably asking the impossible.[15]

Cronbach and his associates (1980) have described the political context of evaluation in ways which make very clear the importance of understanding the politics of evaluation and learning requisite behaviors to work effectively with decision makers. Yet both areas are almost completely neglected in most current training programs. If ways can be found to crowd still more into curricula for evaluator training, it seems prudent to give highest priority to these areas.

Sanders (1981) made distinctions between formal and informal professional development for evaluators, between preservice and inservice professional development, and between training (skills acquisition) and education (acquiring a larger world view). Combining these distinctions into an eight-cell matrix, he suggested means of development in each of the eight resulting categories. Table 3 contains a summary of the ways in which the professional evaluator may gain competence. As one can see, professional development in evaluation is a long-term, complex undertaking that cannot be easily accomplished in a short period of time by a content specialist.

Finally, consider the proposals of Guba and Lincoln (1981) that evaluators be able to empathize with all different points of view and possess the essential traits of good fieldworkers (in the anthropologists' view). This adds an additional dimension to the already complex task of preparing educational evaluators.

6. Professional Status and Rewards

So far the case seems to favor the content-based evaluator. Here we have the best of all possible worlds--a person trained both as an evaluation expert and as a specialist in the relevant content area. But before concluding that this is the ideal, a brief examination of the practicalities of career development and job security is in order.

Content Specialist. There are many first-rate content specialists available in almost every specialization in education and the cognate fields on which education depends. Such experts are familiar figures on site visit evaluation teams, as commentators and sometime critics, and occasionally as co-workers on educational programs or projects. However, very few of these persons aspire to be fulltime evaluators or consent to serve in such roles even if called. This is probably because evaluation would represent for most content specialists a career detour, if not a dead end. Hively and his colleagues outlined the problem for curriculum designers. It seems equally true for evaluators.

> "Curriculum projects frequently enlist the help of numerous scientists and mathematicians. These people may initially volunteer their time and effort as occasional consultants, summer writing-session participants, part-time staff members, etc. But scientists and mathematicians who contribute this kind of time and effort run the danger of losing their professional status and orientation, since the more time they spend away from their own direct, professional pursuits, the less likely they are to stay abreast of their field. The point is not just that subject-matter experts who throw themselves into the effort of curriculum reform may lose out professionally, but that many of them fail to make themselves available at all for fear of falling behind...
>
> "One suggested way of overcoming this problem is to involve not the advanced level expert, but his students...but the same serious flaws remain. Graduate students have presumably embarked on a career in their chosen field, just as their professors. For a graduate student to take time out to study the issues of curriculum design can be threatening to his career, also. The same pressures are likely to build up for him to change his career plans or to be content with second-rate professional status." (Hively, et al., 1973, pp. 52-53).

If this analysis is accurate, then its implications are serious, for it suggests that most of the content specialists willing to become fulltime evaluators are likely to be persons with unimpressive credentials in their own areas of expertise. Obviously there are exceptions to this assertion, but these problems with career socialization and rewards bode ill for the help educational evaluation is likely to draw from content specialists who split their professional time to become part-time evaluation practitioners.

Content-based evaluator. The only difference between the content specialist and the content-based evaluator is that the latter is trained not only in a substantive area but in evaluation methods and techniques as well. If the primary identification of the content-based evaluator is in the substantive field, then taking on evaluation activities threatens primary career interests as much as it does for the content specialist, with the same negative results as mentioned above.

Of course, many content-based evaluators might simply prefer to evaluate within their area of expertise rather than perform other functions in that area. But as long as evaluators restrict their evaluation activities to one content area, their career security seems low. (How many continuing careers can be built solely upon evaluation of programs in linguistics? Is continuing employment probable in a career devoted to evaluating art curricula?) However, if lack of evaluation opportunities within one field causes the content-based evaluator to begin to do evaluations in other areas as well, little

functional difference from the professional evaluator is left, with the exception that the latter is likely to be more experienced in applying evaluation methods to widely differing disciplines and problem areas. Either way, there is little comfort for those who argue for joint training in evaluation and content specialization. The good and gainfully employed content-based evaluator is likely to prove a rare find.

Professional evaluator. Here there is more cause for optimism when one examines the issue of status and rewards. One would expect the most qualified professional evaluators to seek fulltime employment in evaluation, for that is their chosen profession. the only factor which seems to have any potential for deterring talented young methodologists from making a career in evaluation lies in the university context, where there is some uncertainty about how well reward structures serve professional evaluators.[16] Cronbach, et al. (1980) described the job market for evaluators as "capricious," but we do not see the supply of qualified professional evaluators even approaching the market limit at this time.

Summing Up

All things considered, the professional evaluator has had the best of it to this point in the discussion. Except when the evaluator is serving as sole judge in difficult or unique content areas, there seems little need to be a content specialist or a content-base evaluator. Where content specialization is relevant, there is a concern about the independence of an evaluator who holds allegiance in the content area

being evaluated and the possibility that the evaluator feeling comfortable in the area may operate to preclude input from other reference groups. The fact that evaluators' problems are generally set by clients in a variety of content areas points clearly to the futility of attempting to train evaluators in all those areas. Further, the lists of tasks nominated as important in most educational evaluations makes it clear that hiring a reading expert to evaluate a reading project and calling this person an evaluator does not make it so (any more than the reverse would be true). Finally, the content-based evaluator who appears a strong contender because of dual expertise in evaluation and a content area fades when the issues of professional status, rewards and career stability are raised. Surely no one answer is possible as to what combination of evaluation and content expertise is best for evaluations, but the analysis summarized above leads us to suggest that the person trained as a professional evaluator would be the best choice to evaluate most educational enterprises. For this bold assertion to hold up, it is necessary to deal with one final topic--how the professional evaluator can effectively manage those portions of the evaluation where content expertise is needed but not possessed by the evaluator.

Evaluators as Methodologists and Brokers

Recognition that evaluation specialists must elicit input from subject experts is widespread. Scriven (1973, 1974b) called for serious consideration to be given to subject matter experts' opinions of the quality of curriculum materials. Stake (1973) proposed that

evaluators seek out, process, and report opinions of "persons of special qualification," presumably including content specialists. Provus (1973) proposed that subject specialist consultants serve with evaluation specialists as part of a large evaluation team. The Phi Delta Kappa Study Committee (Stufflebeam, et al., 1971) pointed out that subject expertise is not something the evaluator should be expected to possess but rather a resource to be tapped as the "interface role" of the evaluation specialist is fulfilled. Hively, et al., (1973) suggested the use of subject-matter experts as informants and further proposed that the expertise of social scientist be drawn on continually during the process of goal setting. Cronbach, et al., (1980) and the Joint Committee on Standards for Educational Evaluation (1981) likewise recognized the use of teams of experts and the involvement of credible stakeholders for most evaluations. Moreover, advocates of responsive approaches to evaluation (Stake, 1975a,b; Parlett and Hamilton, 1976; Guba and Lincoln, 1981) have developed procedures for evaluator-participant interactions. The concept of stakeholder-based evaluation (Bryk, 1983) has provided guidance for the evaluator in processing information gained from others.

The commonalty in all of these suggestions is that professional evaluators elicit whatever they need (be it value judgments or factual information) from those who are in a better position than they to provide it. There is an apparent divergence in the motivation for seeking outside input between (1) eliciting information from persons because they are in a better position to determine its accuracy (e.g.,

a mathematician checking the accuracy of formulae in an advanced calculus curriculum); and (2) eliciting values or opinions from persons because their expertise or experiences are particularly relevant in reaching a final judgment of worth (e.g., the mathematician judges one calculus curriculum best, despite several errors in formulae, because it presents more important concepts that its competitors and because the errors do not result in serious misconceptions).[17] Content specialization plays an important role in educational evaluations, but it seems neither necessary nor desirable to argue for its inclusion in the preparation of persons who wish to serve as educational evaluators.

FOOTNOTES

[1]The authors' interest in the issues addressed in this paper is reflected in more fragmentary comments on these topics presented in a variety of professional forums during the past decade (Worthen, 1974, 1977, 1978) and Sanders (1979, 1981). Those threads have been drawn together herein in a revisiting of the topic which synthesizes, adds and expands to reflect continuing developments in the field of evaluation.

[2]We will resist altogether the temptation to call the professional evaluator a "content-free evaluator" and the resulting work "content-free evaluation," for fear the term would be a gratuitous contribution to critics of evaluation.

[3]Whether such statements are descriptions or desiderata is not completely clear. They were probably true for most evaluators on national curriculum reform projects flourishing at the time of Scriven's comment (e.g., PSSC Physics), but they seem less descriptive of local curriculum development efforts initiated within the field of education. Locally initiated curricula provide the context in which Tyler (1950) has suggested that curriculum designers should evaluate (or at least assess) the curriculum they produce as part of their professional responsibility. The evaluator and curriculum designer are one and the same here and a greater degree of commitment is difficult to imagine. Nor is that the only setting where curriculum packages and other educational products are evaluated only by persons whose expertise (and often biases as well) lies in the content on which the product is based.

[4]It is true that many educational transactions and outcomes are complex and difficult to understand in the fullest sense of the word. But that is not the same as saying that our level of understanding about those transactions or outcomes is difficult to grasp. Indeed, our understandings of most educational phenomena are no doubt serious underestimates of the real complexity in the phenomena themselves. It is different to say that what we know about a complex phenomenon is relatively little and easy to grasp than to portray what we know as complex because we are aware that our knowledge is only partial.

[5]"Simplicity" is used here to refer to a lack of complication, not to fatuity or silliness.

[6]Two such explanations might be suggested: (1) many simple notions in education have been obfuscated through unclear thinking and the "complexity" may exist only in the mind of the confused; and (2) cries of complexity are often used to warn others away from examining work which is known to be indefensible but is important for political or economic reasons.

[7]Scriven (1974a) and others have argued convincingly against the myth of "value-free science" proposed by writers such as Weber (1949) and Nagel (1961). Agreeing that evaluation and science can in no way be value-free does not require, however, that one expect personal value judgments to shape the conclusions of any form of disciplined inquiry.

[8]Witness the counselor-evaluator who, in evaluating a set of new counseling materials, refuses to accept the objectives of the counselors who are developing the materials and replaces them with an entirely new set. This is not simply evaluating objectives, as Scriven (1973) and Sanders and Cunningham (1973, 1974) would suggest, or transforming the developers' intents into usable objectives as Stake (1973) suggested. This is the content specialist who, because of knowledge of the area, engages in wholesale redirection of the materials, thus misusing the role of evaluator.

[9]The unwary who would use this to argue against careful formulation of objectives should recall two points: (1) Scriven accedes the importance of careful formulation of objectives for the project staff and internal formative evaluation team, and (2) the doctrine of goal-free evaluation which Scriven has suggested as an alternative bypasses not only goals but also the general philosophical statements of rationale often considered by developers as prerequisites for any person evaluating their curriculum. The common failure of educators to differentiate between those unfamiliar with their goals and those who are innocent of the content is relevant here.

[10]This synthesis depends on an approach which Wright (1984) has subsequently defined as the first essential step in competency measurement.

[11]This categorization depends not only on the labeling of the 25 tasks but also on a knowledge of the specific competencies which each task requires. The latter detail appears in the earlier synthesis and is not repeated here.

[12]This does not mean the information is always easy to obtain.

[13]The sources used by Sanders included Sanders (1970), Schalock and Sell (1972), Payne (1974), Worthen (1975), Sanders (1979), and the Joint Committee on Standards for Educational Evaluation (1981).

[14]The former is common in some curriculum development activities, while the latter is more typical in public schools.

[15]The discussion almost begins to be fanciful; in reality, few evaluators are trained as broadly or as well as the ideals outlined above. However, this does not negate the contrast between evaluation and content specialists in this regard.

[16]A discussion of university reward structures as they relate to evaluation appears in Glass and Worthen (1972).

[17] The position that opinions and judgments should be collected from all reference groups directly affected by an education program so those groups can judge and control their own educational systems is not discredited here by omission; it is simply less relevant to the present discussion of subject matter expertise.

Table 1

Need for Evaluation and Content Specialization in 25 Evaluation Tasks

	Task	Task best suited to be done by: Content Specialist	Either	Evaluation Specialist
1.	Obtaining information about a phenomenon to be evaluated	X		
2.	Drawing implications and standards from prior research and practice		X	
3.	Defining the object of the evaluation		X	
4.	Selecting an appropriate inquiry strategy for the evaluation			X
5.	Formulating questions to be answered by the evaluation		X	
6.	Specifying classes of data necessary to answer the evaluation questions		X	
7.	Selecting a design appropriate to answer the questions			X
8.	Identifying the population of interest and selecting appropriate samples			X
9.	Applying the design and controlling threats to validity			X
10.	Identifying the goals of the program to be evaluated	X		
11.	Assessing the value and feasibility of goals			X
12.	Translating goals into measurable objectives			X
13.	Identifying standards or norms for judging worth		X	

14.	Monitoring to detect deviations from design or specified procedures		X
15.	Identifying classes of variables for measurement	X	
16.	Selecting or developing techniques of measurement		X
17.	Assessing the validity of measurement techniques		X
18.	Using appropriate instruments to collect data		X
19.	Choosing appropriate techniques of statistical analysis		X
20.	Using computers and other analysis aids		X
21.	Drawing appropriate conclusions from data analysis		X
22.	Reporting evaluation findings and implications	X	
23.	Making recommendatons based on the evaluation	X	
24.	Providing immediate feedback for use in program management	X	
25.	Obtaining and managing resources to conduct the evaluation	X	

Table 2

Need for Evaluation and Content Specialization in 5 Evaluation Tasks

Task	Task best suited to be done by: Content Specialist	Either	Evaluation Specialist
1. Identifying relevance and applicability of new evaluation approaches			X
2. Describing the evaluation's political context and pressures within that context		X	
3. Dealing with political and interpersonal influences		X	
4. Modifying evaluation approach based on changes in the context or situation			X
5. Maintaining high ethical standards of evaluation			X

Table 3

Professional Development in Evaluation

	PRESERVICE	INSERVICE
TRAINING	1. Courses in Measurement Statistics, Research Methods Evaluation Techniques 2. Supervised applications	1. Workshops, institutes 2. Continuing education - courses, seminars 3. Collegial brown bags 4. Self-instructional packages 5. Use of consultants 6. Quality control within the organization
EDUCATION	1. Courses in Measurement, Statistics, Research Methodology, Evaluation Methodology, Philosophy 2. Internships (arranged) 3. Advanced Seminars in Evaluation	1. Problem-solving seminars 2. Collegial brown bags 3. Conferences 4. Supervisor guidance
TRAINING	1. Peer tutoring 2. Reading (self-motivated) 3. Consultation with others	1. Professional reading 2. Assisting a professional evaluator

EDUCATION	1. Mentor affiliation	1. Professional reading
	2. Modeling	2. Assisting a professional evaluator
	3. Doing/assisting in evaluations	3. Doing evaluations/learning from experience
	4. Extended apprenticeship	4. Networking - sharing experiences

References

Bryk, A. S. (Ed.) (1983). Stakeholder-based evaluation. New Directions for Program Evaluation Series, No. 17. San Francisco, CA: Jossey-Bass.

Cronbach, L. J. (1982). Designing evaluations of educational and social programs. San Francisco, CA: Jossey-Bass.

Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. S., Walker, D. F., & Weiner, S. S. (1980). Toward reform of program evaluation. San Francisco, CA: Jossey-Bass.

Davis, G. G., Scriven, M., & Thomas, S. (1981). The evaluation of composition instruction. Pt. Reyes, CA: Edgepress.

Glass, G. V., & Worthen, B. R. (1971, Fall). Educational evaluation and research: Similarities and differences. Curriculum Theory Network.

Glass, G. V., & Worthen, B. R. (1972). Educational inquiry and the practice of education. In H. D. Schalock & G. R. Sell (Eds.), The Oregon studies in educational research, development, diffusion and evaluation: Vol. III, conceptual frameworks for viewing educational RDD&E. U. S. Office of Educaton Grant No. OEG-0-7-4977. Project No. 0-07001, Monmouth, OR: Teaching Research, Oregon College of Education.

Guba, E. G., & Lincoln, Y. S. (1981). Effective evaluation. San Francisco, CA: Jossey-Bass.

Guttentag, M. (1973). Subjectivity and its use in evaluation research. Evaluation, 1(2), 60-65.

Hively, W., Maxwell, G., Robehl, G., Sension, D., & Lundin, S. (1973). Domain-referenced curriculum evaluation: A technical handbook and a case study from the Minnemast project (CSE Monograph Series in Evaluation, No. 1). Los Angeles: University of California, Center for the Study of Evaluation.

Joint Committee on Standards for Educational Evaluations. (1981). Standards for evaluations of educational programs, projects and materials. New York: McGraw-Hill.

Nagel, E. (1961). The structure of science. New York: Harcourt, Brace and Jovanovich.

Parlett, M., & Hamilton, D. (1976). Evaluation as illumination: A new approach to the study of innovatory programs. In G. V. Glass (Ed.), Evaluation Studies Review Annual (Vol. 1). Beverly Hills, CA: Sage.

Payne, D. A. (1974). Curriculum evaluation. Lexington, MA: D. C. Heath.

Provus, M. M. (1973). Evaluation of ongoing programs in the public school system. In B. R. Worthen & J. R. Sanders, Educational evaluation: Theory and practice. Belmont, CA: Wadsworth.

Sanders, J. R. (1970). Capabilities for evaluators. Bloomington, IN: Indiana University School of Education.

Sanders, J. R. (1979). On the technology and art of evaluation. A review of seven evaluation primers. Evaluation News (12), 11-17.

Sanders, J. R. (1981, November). Evaluation competency and the professional development of evaluators. Invited address at the 1981 joint meeting of the Iowa Educational Research and Evaluation Association and the Midwest Educational Research Association, Des Moines, IA.

Sanders, J. R., & Cunningham, D. J. (197). A structure for formative evaluation in product development. Review of Educational Research, 43, 217-236.

Sanders, J. R., & Cunningham, D. J. (1974). Techniques and procedures for formative valuation. In G. Borich (Ed.), Evaluating Educational Programs and Products. Englewood Cliffs, NJ: Educational Technology Press.

Schalock, H. D., & Sell, G. R. (Eds.). (1972). The Oregon studies in educational research, development, diffusion and evaluation: Vol. III, conceptual frameworks for viewing educational RDD&E. U. S. Office of Education Grant No. OEG-0-70-4977. Project No. 0-0701, Monmouth, OR: Teaching Research, Oregon College of Education.

Scriven, M. (1974a). The exact role of value judgments in science. Berkeley, CA: University of California (mimeo).

Scriven, M. (1974b). Program and product evaluation checklist. In G. D. Borich (Ed.). Evaluating educational programs and products. Englewood Cliffs, NJ: Educational Technology Publications. Also in W. J. Popham (Ed.), Evaluation in education. Berkeley, CA: McCutchen.

Stake, R. E. (1973). The countenance of educational evaluation. In B. R. Worthen & J. R. Sanders, Educational evaluation: Theory and practice, Belmont, CA: Wadsworth.

Stake, R. E. (1975a). Program evaluation, particularly responsive evaluation. (Occasional Paper No. 5). Kalamazoo: Western Michigan University, The Evaluation Center.

Stake, R. E. (1975b). Evaluating the arts in education: A responsive approach. Columbus, OH: Charles E. Merrill.

Stufflebeam, D. L. (1973). Evaluation as enlightenment for decision-making. In B. R. Worthen & J. R. Sanders, Educational evaluation: Theory and practice. Belmont, CA: Wadsworth.

Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. G., Hammond, R. I., Merriman, H. O., & Provus, M. M. (1971). Educational evaluation and decision-making. Itasca, IL: Peacock.

Tyler, R. W. (1950). Basic principles of curriculum instruction. Chicago, IL: University of Chicago Press.

Weber, M. (1949). "Objectivity" in social science and social policy. In E. A. Shils, & H. A. Finch (Eds.) Max Weber on the methodology of the social sciences. Glencoe, IL: Free Press.

Weiss, C. H. (1973). Between the cup and the lip. Evaluation, 1(2), 49-55.

Worthen, B. R. (1974, April). Content specialization and educational evaluation. Presented at the annual meeting of the American Educational Research Association, Chicago, Il.

Worthen, B. R. (1975). Competencies for educational research and evaluation. Educational Researcher, 4, 13-16.

Worthen, B. R. (1977, November). Some thoughts about expertise needed for evaluating curricula. Presented at the Australia Association for Research in Education Conference (Plenary Session), Canberra, Australia.

Worthen, B. R. (1978). Expertise needed in curriculum evaluation. Australian Educational Researcher, 5(2), 5-27.

Worthen, B. R. (1983). An analysis of the evaluation competencies required in three outstanding evaluation studies (Occasional Paper No. 1). Logan, UT: Wasatch Institute for Research and Evaluation.

Worthen, B. R., & Sanders, J. R. (1973). Educational evaluation: Theory and practice. Belmont, CA: Wadsworth.

Wright, B. D. (1984). Essential steps in competency measurement. Competency Research Reports, 1(1), 1-2.